

# The theory of rough paths via one-forms and the extension of an argument of Schwartz to rough differential equations

Terry J. Lyons, Danyu Yang \*

June 15, 2015

## Abstract

We give an overview of the recent approach to the integration of rough paths that reduces the problem to classical Young integration [13]. As an application, we extend an argument of Schwartz [11] to rough differential equations, and prove the existence, uniqueness and continuity of the solution, which is applicable when the driving path takes values in nilpotent Lie group or Butcher group.

## 1 Overview

For each  $p \in [1, \infty)$  Banach introduced a metric for measuring degrees of roughness in paths with values in Banach spaces known as  $p$ -variation. The paths of finite 1-variation are dense in the space of paths of finite  $p$ -variation for each  $p \geq 1$ . Where when  $p = 1$  the paths are weakly differentiable almost surely and they engage with the classical Newtonian calculus for example making sense of line integrals:

$$\int_{t \in [0, T]} \tau_t \otimes d\sigma_t.$$

Young [13] extended the integration so that if  $\tau$  has finite  $q$ -variation and  $\sigma$  is continuous<sup>1</sup> and has finite  $p$ -variation where  $p^{-1} + q^{-1} > 1$  then

$$\int \tau \otimes d\sigma$$

is well defined. In particular, if  $\sigma$  is of finite  $p$ -variation for  $p < 2$  then the integral

$$\int \sigma \otimes d\sigma$$

is meaningfully defined. Young's original definition was directed towards definite integrals. Lyons [6] considered the case of indefinite integrals and the related context of controlled systems of differential equations:

$$dy_t = f(y_t) d\sigma_t, \quad y_0 = a, \tag{1}$$

established the existence and uniqueness of the solution, and also the continuity of the solution in the driving signal. Lyons' integral requires the finite  $p$ -variation of  $\sigma$ , the finite  $\text{Lip}(\gamma)$  norm of  $f$ , and  $p^{-1} + \gamma p^{-1} > 1$ . The methods rely strongly on Young's approach, but a careful examination reveals that the arguments also rely critically on the notion of the Lipschitz function and on the division lemma for them (Proposition 1.26 [8]).

**Lemma 1 (Division Property)** *For Banach spaces  $\mathcal{U}$  and  $\mathcal{W}$ , suppose  $f : \mathcal{U} \rightarrow \mathcal{W}$  is  $\text{Lip}(\gamma)$  for some  $\gamma > 1$ . Then there exists  $h : \mathcal{U} \times \mathcal{U} \rightarrow L(\mathcal{U}, \mathcal{W})$  which is  $\text{Lip}(\gamma - 1)$  such that*

$$f(x) - f(y) = h(x, y)(x - y), \quad \forall x, y \in \mathcal{U},$$

---

\*The authors would like to acknowledge the support of the Oxford-Man Institute and the support provided by ERC advanced grant ESig (agreement no. 291244).

<sup>1</sup>or at least has its jumps in different times to  $\tau$

and for some constant  $C$  depending only on  $\gamma$  and  $\mathcal{U}$ ,

$$\|h\|_{\text{Lip}(\gamma-1)} \leq C \|f\|_{\text{Lip}(\gamma)}.$$

The bound  $p < 2$  becomes an essential part of the thinking if one relies on Young's integral. Both  $p$ -variation paths and  $\text{Lip}(\gamma)$  functions form local algebras, and  $y$  in (1) also has finite  $p$ -variation. From this it is clear that the space of integrals of  $\sigma$ , including all spaces of solutions to differential equations driven by  $\sigma$ , is closed under addition, and the pointwise multiplication is explicitly given by

$$\begin{aligned} \text{for } y_t &= a + \int_{s \in [0,t]} f(y_s) d\sigma_s \text{ and } \hat{y}_t = \hat{a} + \int_{s \in [0,t]} \hat{f}(\hat{y}_s) d\sigma_s, \\ y_t \hat{y}_t &= \int_{s \in [0,t]} \left( f(y_s) \hat{y}_s + y_s \hat{f}(\hat{y}_s) \right) d\sigma_s + a\hat{a}. \end{aligned}$$

This remark is implicit in establishing the existence, uniqueness and continuity theorems since it underpins the operations used in Picard iteration and other approximation strategies. In fact it is easy to show that composition of an integral of  $\sigma$  with a smooth function is also an integral of  $\sigma$  (the chain rule).

In further work [7], Lyons extended the integral of Young to the case  $p \geq 2$ , showed how the notion of bounded variation paths naturally admits a generalization to  $p$ -rough paths for any  $p \in [1, \infty)$ , and established an integral, existence, uniqueness and continuity theorem for differential equations controlled by weak geometric  $p$ -rough paths when  $f$  is  $\text{Lip}(\gamma)$  and  $\gamma > p$ . Young's tricks, the division lemma and the algebraic manipulations of Picard iteration were all important ingredients. The main surprise over the case  $p < 2$  came from the essential nonlinear aspects of the metric imposed on bounded variation functions that allowed the  $p$ -roughness. The space is quite different to that envisaged by Banach.

In this short note we summarize a new approach to the case  $p \geq 2$ , which could be viewed as a proper extension of Lyons' original approach, and is somewhere between the original arguments which emphasized the rough paths and the perspective of Gubinelli which emphasized more the space of possible integrands for a given path that (in his context) are referred to as controlled rough paths. We explain how a clear perspective about a Lipschitz function  $f$  which allows one to (quite simply) reduce the problem of defining a rough line integral

$$\int_{s \in [0,t]} f(\sigma_s) d\sigma_s$$

to the integral of a slowly varying one-form  $t \rightarrow \hat{f}(\sigma_t)$  against a rapidly varying path  $\sigma_t$  in a way that satisfies Young's conditions.

The key understanding comes from repositioning the integral so that  $\sigma$  is a path in a nilpotent group and  $h_t = \hat{f}(\sigma_t)$  is a closed one-form on that group that varies more slowly with time than  $\sigma$ . When looked at in the correct way, Young's strategy applies and

$$\int_{s \in [0,t]} h_s d\sigma_s$$

is well defined. Apart from the clarity this understanding gives, it captures the linearity of the integral against a path in a convenient way, and actually leads to the introduction of the integral of any  $q$ -variation path with values in the closed one-forms against  $\sigma$ . It is not surprising that the class of these integrals is again closed under addition, pointwise multiplication and composition with smooth functions. What is more surprising is that it is (by construction) rich enough to include the original integral

$$\int_{s \in [0,t]} f(\sigma_s) d\sigma_s.$$

As a result, differential equations against rough paths, etc. are easily deduced. It is surprising because  $s \mapsto f(\sigma_s)$  is certainly not in general of finite  $q$ -variation for any  $q$  satisfying

$$\frac{1}{p} + \frac{1}{q} > 1,$$

if  $p \geq 2$ .

The key point is actually rooted in geometry that does not have anything (directly) to do with rough paths but it positions one accurately to do the analysis of rough paths. We need a number of separate ingredients to explain clearly the framework.

**Polynomial functions** A *polynomial function* of degree  $n$  is a globally defined function whose  $(n+1)$ th derivative exists and is identically zero. We intentionally avoid the definition as a power series around a point, and we could choose different reference points and have different representations of the *same* polynomial. More specifically, for Banach spaces  $\mathcal{V}$  and  $\mathcal{U}$ , we say  $p : \mathcal{V} \rightarrow \mathcal{U}$  is a polynomial function of degree (at most)  $n$  if  $D^{n+1}p \equiv 0$ . For any  $y \in \mathcal{V}$ , we can represent  $p$  as a power series around  $y$ :

$$p(x) = \sum_{k=0}^n (D^k p)(y) \frac{(x-y)^{\otimes k}}{k!}, \quad \forall x \in \mathcal{V}, \forall y \in \mathcal{V},$$

but the value of  $p$  does not vary with  $y$ . We would like to emphasize that  $p$  is a function defined on the affine space  $\mathcal{V}$ , it has no natural graded algebraic structure, there is no particular choice of base point associated with it, and there does not exist a translation invariant norm on the space of polynomial functions.

Just as in linear algebra, where one keeps the concept of linear map separated from the matrix one gets after fixing a particular choice of basis, it is conceptually essential that we distinguish the polynomial function as an object from any representation of it via its Taylor series around a chosen point.

For Banach space  $\mathcal{U}$  and integer  $n \geq 0$ , let  $P^{(n)}(\mathcal{U})$  denote the space of polynomial functions of degree  $n$  taking values in  $\mathcal{U}$ .

**Lipschitz functions** By using the polynomial functions (rather than power series), we can shift the classical viewpoint of the Lipschitz function as a function taking values in power series to a function taking values in polynomial functions. This modification gives rise naturally to a way to compare the representations of polynomial functions, and reduces a Lipschitz function to a "slowly-varying" polynomial function. The first author would like to thank Youness Boutaib for sharing his understanding of Lipschitz functions with him.

**Definition 2 (Stein)** Let  $\mathcal{V}$  and  $\mathcal{U}$  be two Banach spaces. For  $\gamma > 0$ , denote  $n := \lfloor \gamma \rfloor$  (the largest integer which is strictly less than  $\gamma$ ). For a closed set  $\mathcal{K}$  in  $\mathcal{V}$ , we say  $f$  is a Lipschitz function of degree  $\gamma$  on  $\mathcal{K}$ , if

$$f : \mathcal{K} \rightarrow P^{(n)}(\mathcal{U}),$$

and for some constant  $M > 0$ ,

$$\sup_{x \in \mathcal{K}} \|f(x)_x\|_\infty + \sup_{x, y \in \mathcal{K}} \max_{j=0,1,\dots,n} \left\| \frac{(D^j(f(x) - f(y)))_x}{\|x - y\|^{\gamma-j}} \right\|_\infty \leq M.$$

Some explanatory points are in order:

1. For  $x \in \mathcal{K}$ ,  $f(x)$  is a polynomial function of degree  $n$ , and we denote by  $f(x)_x$  the degree- $n$  Taylor series of  $f(x)$  around  $x$ . Similarly, for  $j = 0, 1, \dots, n$ ,  $(D^j(f(x) - f(y)))_x$  is a polynomial function of degree  $n - j$  and  $(D^j(f(x) - f(y)))_x$  denotes its degree- $(n - j)$  Taylor series around  $x$ .
2. For each  $x \in \mathcal{K}$ ,  $f(y) \mapsto \|f(y)_x\|_\infty$  is a norm on  $P^{(n)}(\mathcal{U})$ . These norms are equivalent, and if  $\mathcal{K}$  is compact then they are uniformly equivalent.
3. The  $\text{Lip}(\gamma)$  norm  $\|f\|_{\text{Lip}(\gamma)}$  is defined to be the smallest  $M$  satisfying the inequality.

4. Suppose  $\mathcal{N}$  is a neighborhood of  $x$  and  $\mathcal{N} \subseteq \mathcal{K}$ . Then  $F : \mathcal{N} \rightarrow \mathcal{U}$  defined by  $y \mapsto (f(y))(y)$  for  $y \in \mathcal{N}$  is a  $C^\gamma$  function ( $n$  times differentiable with the  $n$ th derivative  $(\gamma - n)$ -Hölder) and  $f(x)$  is the polynomial function that matches  $F$  to degree  $n$  at  $x$  :

$$(D^j (f(x) - F))(x) = 0, \quad j = 0, 1, \dots, n.$$

While in comparison with the notion of  $C^\gamma$  functions, Lipschitz functions make perfect sense even when  $\mathcal{K}$  is of finite cardinality.

5. The space of Lipschitz functions forms an algebra.
6. Whitney's extension theorem was extended by Stein [12] to these generalized Lipschitz functions. He proved that there is a constant  $C_d$  and a linear extension operator so that any  $\text{Lip}(\gamma)$  function  $f$  on a closed set  $\mathcal{K}$  in  $\mathbb{R}^d$  can be extended to a  $\text{Lip}(\gamma)$  function  $g$  on  $\mathbb{R}^d$  where  $\|g\|_{\text{Lip}(\gamma)} \leq C_d \|f\|_{\text{Lip}(\gamma)}$ .

The *crucial* and somewhat counter-intuitive remark associated with Lipschitz functions is the following.

**Remark 3** Suppose  $p$  is a polynomial function of degree  $m$  and  $\gamma > 0$  is a real number. When  $\gamma > m$ ,  $p$  is associated with a constant  $\text{Lip}(\gamma)$  function  $f : \mathcal{K} \rightarrow P^{(m)}(\mathcal{U})$  defined by

$$f(x) := p, \quad \forall x \in \mathcal{K}.$$

When  $\gamma \leq m$ ,  $p$  gives rise to a non-constant  $\text{Lip}(\gamma)$  function

$$f(x)_x(z) = \sum_{l=0}^{\lfloor \gamma \rfloor} (D^l p)(x) \frac{(z-x)^{\otimes l}}{l!}, \quad \forall z \in \mathcal{V}, \quad \forall x \in \mathcal{K},$$

since  $\lfloor \gamma \rfloor < m$ .

**Remark 4** This transformation of polynomials into constant functions in a different function space, and more generally, smooth functions into slowly changing functions, can be seen at the heart of the success of the rough path integral. Rough path integration traditionally integrates a  $\text{Lip}(p + \varepsilon - 1)$  one-form against a (weak geometric)  $p$ -rough path.

**Lifting of polynomial one-forms to closed one-forms** For integer  $n \geq 1$ , the step- $n$  nilpotent Lie group  $G^n$  has a natural graded algebraic structure, and accommodates weak geometric  $p$ -rough paths for  $p < n + 1$ .  $G^1$  is an abelian group which is isomorphic to a Banach space, and fits naturally into the chain  $G^0 = \{e\} \xleftarrow{\pi} G^1 \xleftarrow{\pi} \dots \xleftarrow{\pi} G^n \xleftarrow{\pi} \dots$ . If  $\sigma$  is a path of finite length taking values in  $G^1$ , then there is a natural lift  $\sigma \mapsto \hat{\sigma}$  (the signature of  $\sigma$ ), which takes a path in  $G^1$  into a horizontal path in  $G^n$ .

We have defined polynomial functions and Lipschitz functions. A *polynomial one-form* or a *Lipschitz one-form* is a polynomial function or Lipschitz function taking values in one-forms.

Suppose  $p$  is a polynomial one-form on  $G^1$ , and we would like to lift  $p$  to a one-form  $p^*$  on  $G^n$  so that

$$\int p(\sigma) d\sigma = \int p^*(\hat{\sigma}) d\hat{\sigma}.$$

A simple choice is to let  $p^*$  be the pullback of  $p$  through the projection  $\pi$ . Then the equality holds because  $\sigma = \pi\hat{\sigma}$  and has nothing to do with the fact that  $\hat{\sigma}$  is the "horizontal lift" of  $\sigma$ . Actually, being a "horizontal lift" adds an extra ingredient which we will exploit in a crucial way. If  $\omega$  is any one-form on  $G^n$  which has the horizontal directions in its kernel, then

$$\int p^*(\hat{\sigma}) d\hat{\sigma} = \int (p^* + \omega)(\hat{\sigma}) d\hat{\sigma}.$$

The key point is that we can select  $\omega$  such that  $p^* + \omega$  is a closed one-form, and the selection only depends on  $p$  and not on  $\hat{\sigma}$ .

**Theorem 5** *For  $n \geq 1$  and a polynomial one-form  $p$  of degree  $n - 1$ , there exists a unique one-form  $\omega$  on  $G^n$ , which is orthogonal to the horizontal directions and  $p^* + \omega$  is a closed one-form on  $G^n$ .*

The proof of this theorem is actually not hard: we can give one possible choice of  $\omega$ , and since  $p^* + \omega$  does not depend on  $\hat{\sigma}$ , any two choices must coincide.

While we should specify what we mean by a closed one-form on a group. Roughly speaking, closed one-forms are characterized by zero integral along closed curves, and a one-form on a connected domain is closed if it can be integrated against any continuous path on the domain, and the value of the integral only depends on the end points of the path. A one-form is closed is equivalent to the exact equality between the one-step and two-steps estimates. Integrals often correspond to closed one-forms because of the property  $\int_{[s,t]} = \int_{[s,u]} + \int_{[u,t]}$ , and this property is actually behind the fact that the lifted polynomial one-form is closed. In term of mathematical expression, we say  $\beta$  on group  $\mathcal{G}$  taking values in another algebra is closed (or cocyclic), if

$$\beta(a, b) \beta(ab, c) = \beta(a, bc), \forall a, b, c \in \mathcal{G}.$$

By lifting a path to a horizontal path and a polynomial one-form to a closed one-form on the nilpotent Lie group, we replace a general integral by the integral of a closed one-form. The integral of a closed one-form has the nice property that it does not depend on the fine structure of the path but only on its end points. In particular, the integral makes sense for any continuous path and has no (further) regularity assumption.

**Integrating slowly-varying closed one-forms** Since the integral of a closed one-form against any continuous path is well-defined, we could weaken the requirement on the one-form and strengthen the regularity assumption on the path in such a way that the integral still makes sense. For example, in the case of classical integral, we can integrate a constant one-form against any continuous path because constant one-forms are closed. Then if we weaken the requirement on the one-form and strengthen the requirement on the path in such a way that their regularities "compensate" each other, then the integral still makes sense as Young integral [13]. In the case of Young integral, we actually vary the constant one-form with time and get a path taking values in constant one-forms, which is more clearly seen in the proof of the existence of the integral where we keep comparing the constant one-forms from different times based on their effect on the future increment of the driving path.

Constant one-form on Banach space is just a special example of closed one-forms. More generally, suppose we have a family of closed one-forms on a differential manifold or on a topological group. For a given path taking values in the manifold or group, if the closed one-form varies with time in such a way that the one-form and the path have compensated Young regularities, then the integral should still makes sense.

As we mentioned above, a Lipschitz one-form could be viewed as a slowly-varying polynomial one-form, and that there exists a canonical lift of a polynomial one-form to a closed one-form on the nilpotent Lie group. Hence we can lift a Lipschitz one-form to a slowly-varying closed one-form on the nilpotent Lie group. More specifically, suppose  $\alpha$  is a Lipschitz one-form on  $G^1$ . Then based on our argument above,  $\alpha$  can be viewed as a slowly-varying polynomial one-form. Suppose  $\sigma$  is an underlying reference path. Then the evolution of  $\sigma$  gives a natural order (or say time), and  $\alpha$  along  $\sigma$  is a "slowly-time-varying" polynomial one-form with each  $\alpha_{\sigma_t}$  a polynomial one-form. If we denote by  $\hat{\sigma}_t \in G^n$  the horizontal lift of the path  $\sigma_t \in G^1$  and denote by  $\beta_{\hat{\sigma}_t}$  the closed one-form lift of the polynomial one-form  $\alpha_{\sigma_t}$ , then we can rewrite the integral of a Lipschitz one-form against  $\sigma$  as the integral of a time-varying closed one-form against  $\hat{\sigma}$  :

$$\int \alpha(\sigma_t) d\sigma_t = \int \alpha_{\sigma_t}(\sigma_t) d\sigma_t = \int \beta_{\hat{\sigma}_t}(\hat{\sigma}_t) d\hat{\sigma}_t.$$

When  $\sigma$  is of finite length, this algebraic/geometrical reformulation seems unnecessary. The point is that for general path  $\hat{\sigma}$  of finite  $p$ -variation taking values in  $G^{[p]}$ , the integral  $\int \beta_{\hat{\sigma}}(\hat{\sigma}) d\hat{\sigma}$  still makes sense (the rough integral) while the classical Riemann sum integral  $\int \alpha(\sigma) d\sigma$  does not have a proper meaning.

**Theorem 6** Suppose  $\alpha$  is a  $\text{Lip}(p + \epsilon - 1)$  one-form for some  $\epsilon > 0$ . Then there exists  $\beta$  taking values in closed (or say cocyclic) one-forms on  $G^{[p]}$ , such that for any  $\sigma_t \in G^1$  of finite length with horizontal lift  $\hat{\sigma}_t \in G^{[p]}$ , we have

$$\int \alpha(\sigma_t) d\sigma_t = \int \beta_{\hat{\sigma}_t}(\hat{\sigma}_t) d\hat{\sigma}_t,$$

Moreover, the integral  $\int \beta_{\hat{\sigma}_t}(\hat{\sigma}_t) d\hat{\sigma}_t$  is well-defined for any continuous path  $\hat{\sigma}$  of finite  $p$ -variation taking values in  $G^{[p]}$  and the integral is continuous with respect to  $\hat{\sigma}$  in  $p$ -variation metric.

**Conclusion** Based on our formulation, to make sense of the rough integral, all we need is the compensated Young regularity between two dual paths: one takes values in the group and the other takes values in the closed (cocyclic) one-forms on the group. By viewing the Lipschitz functions as slowly-varying polynomial functions and by lifting the polynomial one-forms to closed one-forms, we encapsulate the nonlinearity of the integral to the structure of the group and to the closed one-forms on the group so that the idea behind the generalized integral is clearer and bears a similar form to the linear Young integral.

## 2 Definitions and Properties

Suppose  $\mathcal{U}, \mathcal{V}$  and  $\mathcal{W}$  are Banach spaces and  $p \geq 1$  a real number. We restate the definition of the cocyclic one-form and the dominated path as in [9].

Suppose  $\mathcal{A}$  and  $\mathcal{B}$  are Banach algebras and  $\mathcal{G}$  is a topological group in  $\mathcal{A}$ . We denote by  $L(\mathcal{A}, \mathcal{B})$  the set of continuous linear mappings from  $\mathcal{A}$  to  $\mathcal{B}$ , and we denote by  $C(\mathcal{G}, L(\mathcal{A}, \mathcal{B}))$  the set of continuous mappings from  $\mathcal{G}$  to  $L(\mathcal{A}, \mathcal{B})$ .

**Definition 7 (Cocyclic One-Form)** We say  $\beta \in C(\mathcal{G}, L(\mathcal{A}, \mathcal{B}))$  is a cocyclic one-form, if there exists a topological group  $\mathcal{H}$  in  $\mathcal{B}$  such that  $\beta(a, b) \in \mathcal{H}$  for all  $a, b \in \mathcal{G}$  and

$$\beta(a, b) \beta(ab, c) = \beta(a, bc), \quad \forall a, b, c \in \mathcal{G}.$$

We denote the set of cocyclic one-forms by  $B(\mathcal{G}, \mathcal{H})$  (or  $B(\mathcal{G})$ ).

Since a Banach space  $\mathcal{U}$  is canonically embedded in the Banach algebra  $\{(c, u) | c \in \mathbb{R}, u \in \mathcal{U}\}$  with multiplication  $(c, u)(r, v) = (cr, ru + cv)$ , we denote by  $B(\mathcal{G}, \mathcal{U})$  the set of cocyclic one-forms taking values in  $\mathcal{U}$  satisfying  $\beta(a, b) + \beta(ab, c) = \beta(a, bc)$  for all  $a, b, c \in \mathcal{G}$ .

For  $p \geq 1$ , we denote by  $[p]$  the integer part of  $p$ . As in [9], we equip the tensor powers of  $\mathcal{V}$  with admissible norms and assume  $T^{([p])}(\mathcal{V}) = \mathbb{R} \oplus \mathcal{V} \oplus \dots \oplus \mathcal{V}^{\otimes [p]}$  is a graded Banach algebra equipped with the norm  $\|\cdot\| := \sum_{k=0}^{[p]} \|\pi_k(\cdot)\|$  ( $\pi_k$  denotes the projection to  $\mathcal{V}^{\otimes k}$ ), and the multiplication on  $T^{([p])}(\mathcal{V})$  is induced by a finite family of linear projective mappings denoted by  $\mathcal{P}_{[p]}$ ;  $\mathcal{G}_{[p]}$  is a closed topological group in  $T^{([p])}(\mathcal{V})$  whose linear span is  $T^{([p])}(\mathcal{V})$  and whose projection to  $\mathbb{R}$  is 1.

When  $\mathcal{G}_{[p]}$  is the nilpotent Lie group over  $\mathcal{V}$ ,  $\mathcal{P}_{[p]} = \{\pi_k\}_{k=0}^{[p]}$  with  $\pi_k(ab) = \sum_{j=0}^k \pi_j(a) \otimes \pi_{k-j}(b)$  for  $k = 0, 1, \dots, [p]$  and for  $a, b \in T^{([p])}(\mathcal{V})$ . When  $\mathcal{G}_{[p]}$  is the Butcher group over  $\mathbb{R}^d$ ,  $\mathcal{P}_{[p]}$  is the set of labelled forests of degree less or equal to  $[p]$  and  $\sigma(ab) = \sum_c P^c(\sigma)(a) R^c(\sigma)(b)$  for  $\sigma \in \mathcal{P}_{[p]}$  and for  $a, b \in T^{([p])}(\mathbb{R}^d)$  where the sum is over all admissible cuts of the forest  $\sigma$ . For more details see [10, 7, 1, 2, 4].

We equip  $\mathcal{G}_{[p]}$  with the norm  $|\cdot| := \sum_{k=1}^{[p]} \|\pi_k(\cdot)\|^{\frac{1}{k}}$  and define the  $p$ -variation of a continuous path  $g : [0, T] \rightarrow \mathcal{G}_{[p]}$  by

$$\|g\|_{p\text{-var}, [0, T]} := \sup_{D, D \subset [0, T]} \left( \sum_{k, t_k \in D} |g_{t_k}^{-1} g_{t_{k+1}}|^p \right)^{\frac{1}{p}}.$$

We denote by  $C^{p\text{-var}}([0, T], \mathcal{G}_{[p]})$  the set of continuous paths of finite  $p$ -variation on  $[0, T]$  taking values in  $\mathcal{G}_{[p]}$ . (The exact form of norm on  $\mathcal{G}_{[p]}$  is not important, and the integral can be defined as long as the norm on the group and the norm on the one-form "compensate" each other.)

For  $\alpha \in L(T^{[p]}(\mathcal{V}), \mathcal{U})$ , we denote

$$\|\alpha(\cdot)\| := \sup_{v \in T^{[p]}(\mathcal{V}), \|v\|=1} \|\alpha(v)\|, \quad \|\alpha(\cdot)\|_k := \sup_{v \in \mathcal{V}^{\otimes k}, \|v\|=1} \|\alpha(v)\|, \quad k = 1, 2, \dots, [p].$$

We say  $\omega : \{(s, t) \mid 0 \leq s \leq t \leq T\} \rightarrow \overline{\mathbb{R}^+}$  is a control, if  $\omega$  is continuous, non-negative, vanishes on the diagonal and satisfies  $\omega(s, u) + \omega(u, t) \leq \omega(s, t)$  for  $0 \leq s \leq u \leq t \leq T$ . As in [9], for  $g \in C([0, T], \mathcal{G}_{[p]})$  and  $\beta : [0, T] \rightarrow B(\mathcal{G}_{[p]}, \mathcal{U})$ , if the limit exists

$$\lim_{|D| \rightarrow 0, D = \{t_k\}_{k=0}^n \subset [0, T]} \beta_0(g_0, g_{0,t_1}) \beta_{t_1}(g_{t_1}, g_{t_1,t_2}) \cdots \beta_{t_{n-1}}(g_{t_{n-1}}, g_{t_{n-1},T}) \quad \text{with } g_{s,t} := g_s^{-1} g_t,$$

then we define the limit to be the integral  $\int_0^T \beta_u(g_u) dg_u$ .

**Definition 8 (Dominated Path)** For  $g \in C^{p-var}([0, T], \mathcal{G}_{[p]})$  and Banach space  $\mathcal{U}$ , we say a continuous path  $\rho : [0, T] \rightarrow \mathcal{U}$  is dominated by  $g$ , if there exists  $\beta : [0, T] \rightarrow B(\mathcal{G}_{[p]}, \mathcal{U})$  which satisfies, for some  $M > 0$ , control  $\omega$  and  $\theta > 1$ ,

$$\begin{aligned} \|\beta_t(g_t, \cdot)\| &\leq M, \quad \forall t \in [0, T], \\ \|(\beta_t - \beta_s)(g_t, \cdot)\|_k &\leq \omega(s, t)^{\theta - \frac{k}{p}}, \quad \forall 0 \leq s \leq t \leq T, \quad k = 1, 2, \dots, [p], \end{aligned}$$

such that  $\rho_t = \rho_0 + \int_0^t \beta_u(g_u) dg_u$  for  $t \in [0, T]$ .

Based on the definition of dominated paths, we introduce an operator norm on the space of one-forms to quantify the convergence of one-forms (associated with Picard iterations).

For  $g \in C^{p-var}([0, T], \mathcal{G}_{[p]})$  and control  $\omega$ , we say  $g$  is controlled by  $\omega$  if  $\|g\|_{p-var, [s, t]}^p \leq \omega(s, t)$  for all  $s < t$ .

**Definition 9 (Operator Norm)** For  $g \in C^{p-var}([0, T], \mathcal{G}_{[p]})$  controlled by  $\omega$  and  $\beta : [0, T] \rightarrow B(\mathcal{G}_{[p]}, \mathcal{U})$ , we define, for  $\gamma > 1$ ,

$$\|\beta\|_\gamma := \sup_{t \in [0, T]} \|\beta_t(g_t, \cdot)\| + \max_{k=1, \dots, \lfloor \gamma \rfloor} \sup_{0 \leq s \leq t \leq T} \frac{\|(\beta_t - \beta_s)(g_t, \cdot)\|_k}{\omega(s, t)^{\frac{\gamma-k}{p}}}.$$

Suppose  $\|\beta\|_\gamma < \infty$ . When  $\gamma$  increases, the integrability of  $\beta$  increases. In the extreme case that  $\gamma$  tends to infinity,  $\beta$  is compelled to be a constant cocyclic one-form, so is integrable against any continuous path. If  $\gamma > p - 1$  and if there exists  $\sigma : [0, T] \rightarrow \mathcal{U}$  such that  $\|\sigma_t - \sigma_s - \beta_s(g_s, g_{s,t})\| \leq C \|g\|_{p-var, [s, t]}^\gamma$  for all  $s < t$ , then  $\sigma$  is a weakly controlled path introduced by Gubinelli [3]. When  $\gamma > p$ ,  $\beta$  is integrable against  $g$  and  $t \mapsto \int_0^t \beta(g_u) dg_u$  is a dominated path.

**Definition 10** Suppose there exists a mapping  $\mathcal{I}' \in L(T^{([p])}(\mathcal{V}), T^{([p])}(\mathcal{V})^{\otimes 2})$  which satisfies

$$\mathcal{I}'(1) = \mathcal{I}'(\mathcal{V}) = 0, \quad \mathcal{I}'(\mathcal{V}^{\otimes k}) \subseteq \mathcal{V}^{\otimes(k-1)} \otimes \mathcal{V}, \quad k = 2, \dots, [p],$$

and (with  $1'_{n,2}$  denoting the projection of  $T^{([p])}(\mathcal{V})^{\otimes 2}$  to  $\sum_{k=1}^{[p]-1} \mathcal{V}^{\otimes k} \otimes \mathcal{V}$ )

$$\mathcal{I}'(ab) = \mathcal{I}'(a) + 1'_{n,2}((a \otimes a) \mathcal{I}'(b)) + 1'_{n,2}((a-1) \otimes (a(b-1))), \quad \forall a, b \in \mathcal{G}_{[p]}.$$

Due to the special form of the dominated paths in Picard iterations, we only need the mapping  $\mathcal{I}'$  (instead of  $\mathcal{I}$  as in [9]) for the recursive integrals to make sense. Roughly speaking, the mapping  $\mathcal{I}$  is used to define the iterated integral of two dominated (controlled) paths, and corresponds to a universal continuous linear mapping which has the "formal" expression:

$$\mathcal{I}(a) = \int_0^T (g_{0,u} - 1) \otimes \delta g_{0,u}, \quad g \in C([0, T], \mathcal{G}_{[p]}), \quad a = g_{0,T}, \quad \forall a \in \mathcal{G}_{[p]}.$$

The mapping  $\mathcal{I}'$  encodes part of the information of  $\mathcal{I}$ , is used to define the integral of a dominated (controlled) path against the first level of the given group-valued path, and corresponds to a universal continuous linear mapping with the formal expression:

$$\mathcal{I}'(a) = \int_0^T (g_{0,u} - 1) \otimes \delta x_u, \quad x := \pi_1(g), \quad g \in C([0, T], \mathcal{G}_{[p]}), \quad a = g_{0,T}, \quad \forall a \in \mathcal{G}_{[p]}.$$

In particular,  $\mathcal{I}'$  is well-defined for degree- $[p]$  nilpotent Lie group and degree- $[p]$  Butcher group for any  $p \geq 1$  (see [9] for more explanation).

The lemma below proves that one can integrate a weakly controlled path [3, 4] and get a dominated path. We made the dependence of the coefficients explicit to suit the special needs of our proof.

**Lemma 11** *Suppose  $g \in C^{p-var}([0, T], \mathcal{G}_{[p]})$  is controlled by  $\omega$ ,  $\beta : [0, T] \rightarrow B(\mathcal{G}_{[p]}, L(\mathcal{V}, \mathcal{W}))$  satisfies*

$$\|\beta\|_\gamma < \infty \text{ for some } \gamma \in (p-1, [p]],$$

*and there exists  $\varphi : [0, T] \rightarrow L(\mathcal{V}, \mathcal{W})$  which satisfies for some  $M > 0$ ,*

$$\|\varphi_t - \varphi_s - \beta_s(g_s, g_{s,t})\| \leq M \|\beta\|_\gamma \omega(s, t)^{\frac{2}{p}}, \quad \forall 0 \leq s < t \leq T. \quad (2)$$

*If we define  $\eta : [0, T] \rightarrow B(\mathcal{G}_{[p]}, \mathcal{W})$  by*

$$\eta_t(a, b) := \varphi_t \pi_1(g_t^{-1} a (b - 1)) + \beta_t(g_t, \cdot) \pi_1(\cdot) \mathcal{I}'(g_t^{-1} a (b - 1)), \quad \forall a, b \in \mathcal{G}_{[p]}, \quad \forall t \in [0, T],$$

*then for some structural constant  $c$  (depending on the mapping  $\mathcal{I}'$ ),*

$$\sup_{0 \leq t \leq T} \|\eta_t(g_t, \cdot)\| \leq \sup_{0 \leq t \leq T} \|\varphi_t\| + c \|\beta\|_\gamma,$$

*and there exists a constant  $C = C(M, p, \omega(0, T))$  such that*

$$\|(\eta_t - \eta_s)(g_t, \cdot)\|_k \leq C \|\beta\|_\gamma \omega(s, t)^{\frac{\gamma+1-k}{p}}, \quad \forall s < t, \quad k = 1, 2, \dots, [p].$$

*As a consequence,  $\|\eta\|_{\gamma+1} < \infty$  and  $t \mapsto \int_0^t \eta_u(g_u) dg_u$  is a dominated path.*

**Proof.** It is clear that for some constant  $c$  depending on  $\mathcal{I}'$ ,

$$\|\eta_t(g_t, \cdot)\| \leq \|\varphi_t\| + c \|\beta_t(g_t, \cdot)\| \leq \sup_{0 \leq t \leq T} \|\varphi_t\| + c \|\beta\|_\gamma, \quad \forall t \in [0, T].$$

For  $s < t$  and  $v \in \mathbb{R} \oplus \mathcal{V} \oplus \dots \oplus \mathcal{V}^{\otimes [p]}$  (calculation or based on the proof in [9]), we have

$$\begin{aligned} (\eta_t - \eta_s)(g_t, v) &= (\varphi_t - \varphi_s - \beta_s(g_s, g_{s,t})) \pi_1(v) + (\beta_t - \beta_s)(g_t, \cdot) \pi_1(\cdot) \mathcal{I}'(v) \\ &\quad + \sum_{\sigma \in \mathcal{P}_{[p]}, |\sigma|=[p]} \beta_s(g_s, \sigma(g_{s,t})) \pi_1(v) \\ &\quad + \sum_{k=2}^{[p]} \sum_{\sigma \in \mathcal{P}_{[p]}, |\sigma| \geq [p]+1-k} \beta_s(g_s, \sigma(g_{s,t}, \cdot)) \pi_1(\cdot) \mathcal{I}'(\pi_k(v)). \end{aligned} \quad (3)$$

Since  $\|\beta\|_\gamma < \infty$  and  $\mathcal{I}'(\mathcal{V}^{\otimes k}) \subseteq \mathcal{V}^{\otimes(k-1)} \otimes \mathcal{V}$ ,  $k = 2, \dots, [p]$ , we have, for some structural constant  $C$  depending on the norm of the mapping  $\mathcal{I}'$ ,

$$\sup_{v \in \mathcal{V}^{\otimes k}, \|v\|=1} \|(\beta_t - \beta_s)(g_t, \cdot) \pi_1(\cdot) \mathcal{I}'(v)\| \leq C \|\beta\|_\gamma \omega(s, t)^{\frac{\gamma+1-k}{p}}, \quad k = 1, 2, \dots, [p].$$

Moreover, for  $s < t$ ,

$$\begin{aligned} \|\beta_s(g_s, \sigma(g_{s,t}))\| &\leq \|\beta\|_\gamma \omega(s, t)^{\frac{[p]}{p}}, \quad \forall \sigma \in \mathcal{P}_{[p]}, \quad |\sigma| = [p], \\ \|\beta_s(g_s, \sigma(g_{s,t}, \cdot)) \pi_1(\cdot) \mathcal{I}'(\pi_k(\cdot))\| &\leq \|\beta\|_\gamma (1 \vee \omega(0, T)) \omega(s, t)^{\frac{[p]+1-k}{p}}, \quad \forall \sigma \in \mathcal{P}_{[p]}, \quad |\sigma| \geq [p] + 1 - k. \end{aligned}$$



Hence, since  $\gamma \leq [p]$ , combined with (3) and (2), for some  $C = C(M, p, \omega(0, T))$ , we have

$$\|(\eta_t - \eta_s)(g_t, \cdot)\|_k \leq C \|\beta\|_\gamma \omega(s, t)^{\frac{\gamma+1-k}{p}}, \quad \forall s < t, k = 1, 2, \dots, [p].$$

■

For  $\gamma \geq 1$ ,  $[\gamma]$  denotes the largest integer which is strictly less than  $\gamma$ . For  $\sigma_i \in \mathcal{P}_{[p]}$ ,  $i = 1, \dots, l$ ,  $|\sigma_1| + \dots + |\sigma_l| \leq [p]$ , we denote by  $\sigma_1 * \dots * \sigma_l$  the continuous linear mapping from  $\mathcal{V}^{\otimes(|\sigma_1| + \dots + |\sigma_l|)}$  to  $\mathcal{V}^{\otimes|\sigma_1|} \otimes \dots \otimes \mathcal{V}^{\otimes|\sigma_l|}$  satisfying  $(\sigma_1 * \dots * \sigma_l)(a) = \sigma_1(a) \otimes \dots \otimes \sigma_l(a)$  for all  $a \in \mathcal{G}_{[p]}$  (see [9] for more details).

**Definition 12** ( $\beta(f(\rho))$ ) *Let  $\rho_\cdot = \rho_0 + \int_0^\cdot \beta(g) dg : [0, T] \rightarrow \mathcal{U}$  be a dominated path and  $f : \mathcal{U} \rightarrow \mathcal{W}$  be a  $\text{Lip}(\gamma)$  function for some  $\gamma > p - 1$ . We define  $\beta(f(\rho)) : [0, T] \rightarrow B(\mathcal{G}_{[p]}, \mathcal{W})$  by, for  $a, b \in \mathcal{G}_{[p]}$  and  $s \in [0, T]$ ,*

$$\beta(f(\rho))_s(a, b) = \sum_{l=1}^{[\gamma]} \frac{1}{l!} (D^l f)(f(\rho_s)) \beta_s(g_s, \cdot)^{\otimes l} \sum_{\sigma_i \in \mathcal{P}_{[p]}, |\sigma_1| + \dots + |\sigma_l| \leq [p]} (\sigma_1 * \dots * \sigma_l)(g_s^{-1} a (b - 1)).$$

**Definition 13 (Integral)** *Suppose  $\rho : [0, T] \rightarrow \mathcal{U}$  is a path dominated by  $g \in C^{p-\text{var}}([0, T], \mathcal{G}_{[p]})$  and  $f : \mathcal{U} \rightarrow L(\mathcal{V}, \mathcal{W})$  is a  $\text{Lip}(\gamma)$  function for some  $\gamma > p - 1$ . With  $\beta(f(\rho))$  in Definition 12, if we define  $\beta : [0, T] \rightarrow B(\mathcal{G}_{[p]}, \mathcal{W})$  by*

$$\beta_s(a, b) = f(\rho_s) \pi_1(g_s^{-1} a (b - 1)) + \beta(f(\rho))_s(g_s, \cdot) \otimes \pi_1(\cdot) \mathcal{I}'(g_s^{-1} a (b - 1)), \quad \forall a, b \in \mathcal{G}_{[p]}, \forall s, \quad (4)$$

then  $\beta$  is integrable against  $g$  and we define the integral  $\int f(\rho) dx : [0, T] \rightarrow \mathcal{W}$  by

$$\int_0^t f(\rho_u) dx_u := \int_0^t \beta_u(g_u) dg_u, \quad \forall t \in [0, T].$$

That  $\beta$  is integrable against  $g$  follows from Lemma 11. When  $\mathcal{G}_{[p]}$  is the nilpotent Lie group, the integral coincides with the first level of the rough integral in [7]. When  $\mathcal{G}_{[p]}$  is the Butcher group the integral coincides with the integral in [4].

**Definition 14 (Solution)** *For  $\gamma + 1 > p \geq 1$ , suppose  $g \in C^{p-\text{var}}([0, T], \mathcal{G}_{[p]})$  and  $f : \mathcal{U} \rightarrow L(\mathcal{V}, \mathcal{U})$  is a  $\text{Lip}(\gamma)$  function. We say  $y$  is a solution to the rough differential equation (with  $x := \pi_1(g)$ )*

$$dy = f(y) dx, \quad y_0 = \xi \in \mathcal{U}, \quad (5)$$

if  $y$  is a path dominated by  $g$ , and  $y_\cdot = \xi + \int_0^\cdot f(y_u) dx_u$  with the integral defined in Definition 13.

Since dominated paths are defined through integrable one-forms, instead of formulating the fixed-point problem in the space of paths as in Definition 14, we could also formulate the fixed-point problem in the space of integrable one-forms, and  $y$  is called a solution to (5) if the one-form associated with  $y$  is a fixed point of the mapping  $\beta \mapsto \hat{\beta}$  where  $\hat{\beta}$  is the one-form associated with  $\int f(y) dx$ .

### 3 Existence, Uniqueness and Continuity of the Solution

Schwartz gave a beautiful proof in [11] of the convergence of the series of Picard iterations for SDEs. Instead of working with contraction mapping on small intervals and pasting the local solutions together, he used the iterative expression of the differences between the  $n$ th and  $(n+1)$ th Picard iterations and proved that the sequence of differences decay factorially on the whole interval. Put in the simplest form, his argument can be summarized as follows. Suppose  $f$  is  $\text{Lip}(1)$  and consider the SDE:

$$dX_t = f(X_t) dB_t, \quad X_0 = \xi.$$

We define the series of Picard iterations by  $X_t^{n+1} = \xi + \int_0^t f(X_u^n) dB_u$  with  $X_t^0 \equiv \xi$ . Then by using Itô's isometry and the Lipschitz property of  $f$ , we have

$$E \left( |X_t^{n+1} - X_t^n|^2 \right) = E \int_0^t |f(X_u^n) - f(X_u^{n-1})|^2 du \leq \|f\|_{\text{Lip}(1)}^2 \int_0^t E \left( |X_u^n - X_u^{n-1}|^2 \right) du.$$

By iterating this process, we obtain a factorial decay and the global convergence of the Picard series.

We will try to extend his argument to RDEs. However, there are several points to pay attention to: generally,  $\text{Lip}(1)$  is insufficient for rough integral to be well-defined and it is illegitimate to take modulus inside the rough integral; there is no  $L^2$  space and no Itô's isometry for general rough paths, so the factorial decay can not be obtained in a similar way. We will rely critically on the Division Property of Lipschitz functions, and rely critically on the factorial decay of the Signature of a rough path [7]. In particular, we prove that the one-forms associated with the differences between the  $n$ th and  $(n+1)$ th Picard iterations decay factorially in operator norm as  $n$  tends to infinity on the whole interval. As a consequence, the one-forms associated with the Picard iterations converge in operator norm, which implies the convergence of the Picard iterations and the convergence of their group-valued enhancements. By using the factorial decay of the iterated integrals, we can prove the solution is unique. The continuity of the solution with respect to the driving noise follows from the uniform convergence of the Picard iterations for the rough differential equations whose driving rough paths are uniformly bounded in  $p$ -variation.

Let  $\mathcal{U}$  and  $\mathcal{V}$  be two Banach spaces.

**Definition 15 (Picard Iterations)** For  $\gamma + 1 > p \geq 1$ , suppose  $g \in C^{p\text{-var}}([0, T], \mathcal{G}_{[p]})$ ,  $f : \mathcal{U} \rightarrow L(\mathcal{V}, \mathcal{U})$  is a  $\text{Lip}(\gamma)$  function and  $\xi \in \mathcal{U}$ . We define the series of Picard iterations associated with the rough differential equation  $dy = f(y) dx$ ,  $y_0 = \xi$ , by

$$y_t^n := \xi + \int_0^t f(y_u^{n-1}) dx_u, \forall t \in [0, T], \text{ with } y_t^0 \equiv \xi.$$

**Definition 16** We define  $\zeta^n : [0, T] \rightarrow B(\mathcal{G}_{[p]}, \mathcal{U})$ ,  $n \geq 1$ , by

$$\zeta_s^n(a, b) = f(y_s^{n-1}) \pi_1(g_s^{-1} a (b - 1)) + \beta(f(y_s^{n-1}))_s(g_s, \cdot) \pi_1(\cdot) \mathcal{I}'(g_s^{-1} a (b - 1)), \forall a, b \in \mathcal{G}_{[p]}, \forall s,$$

where  $\beta(f(y^{n-1}))$  is defined in term of  $y^{n-1} = \xi + \int_0^\cdot \zeta^{n-1}(g) dg$  as in Definition 12 with  $\zeta^0 \equiv 0$ .

Then based on the definition of the integral in Definition 13,  $y^n = \xi + \int_0^\cdot \zeta^n(g) dg$ ,  $n \geq 0$ , and  $\{y^n\}_{n=0}^\infty$  are paths dominated by  $g$ .

**Lemma 17** Suppose  $g \in C^{p\text{-var}}([0, T], \mathcal{G}_{[p]})$  is controlled by  $\omega$ , and  $f : \mathcal{U} \rightarrow L(\mathcal{V}, \mathcal{U})$  is a  $\text{Lip}(\gamma)$  function for some  $\gamma \in (p-1, [p]]$ . Then there exists a constant  $C = C(p, \gamma, \|f\|_{\text{Lip}(\gamma)}, \omega(0, T))$  such that

$$\sup_{n \geq 0} \|\zeta^n\|_{\gamma+1} \leq C.$$

**Proof.** We first suppose  $\omega(0, T) \leq 1$ , and prove that there exists  $\lambda_{p, \gamma} > 0$  which only depends on  $p$  and  $\gamma$  such that when  $\|f\|_{\text{Lip}(\gamma)} \leq \lambda_{p, \gamma}$  we have  $\sup_{n \geq 0} \|\zeta^n\|_{\gamma+1} \leq 2\lambda_{p, \gamma}$ . We prove it by using mathematical induction. Suppose for some constant  $\lambda_n \in (0, 1)$ ,

$$\|\zeta^n\|_{\gamma+1} \leq \lambda_n,$$

which holds when  $n = 0$  since  $\zeta^0 \equiv 0$ . We want to prove that there exists a constant  $C_{p, \gamma} \geq 1$  such that

$$\|\zeta^{n+1}\|_{\gamma+1} \leq \|f\|_{\text{Lip}(\gamma)} (1 + C_{p, \gamma} \lambda_n) := \lambda (1 + C_{p, \gamma} \lambda_n).$$

Then when  $\lambda \in (0, (2C_{p, \gamma})^{-1})$ , if  $\lambda_n \leq \lambda / (1 - C_{p, \gamma} \lambda)$  then  $\lambda (1 + C_{p, \gamma} \lambda_n) \leq \lambda / (1 - C_{p, \gamma} \lambda)$ . Since  $\lambda_0 = 0 \leq \lambda / (1 - C_{p, \gamma} \lambda)$ , we have  $\lambda_n \leq \lambda / (1 - C_{p, \gamma} \lambda) \leq 2\lambda$  for all  $n \geq 0$ . It can be checked that  $\zeta^{n+1}$  is linear with respect to scalar multiplication of  $f$ , so we assume  $\|f\|_{\text{Lip}(\gamma)} = 1$ , and want to prove

$$\|\zeta^{n+1}\|_{\gamma+1} \leq 1 + C_{p, \gamma} \lambda_n \text{ when } \|\zeta^n\|_{\gamma+1} \leq \lambda_n. \quad (6)$$

By following similar proof as in [9] of the stability of dominated paths under composition with regular functions and by using  $\|f\|_{\text{Lip}(\gamma)} = 1$ ,  $\omega(0, T) \leq 1$  and  $\|\zeta^n\|_{\gamma+1} \leq \lambda_n \in (0, 1)$ , we have that there exists  $C_{p,\gamma} > 0$  such that for any  $s < t$ ,

$$\begin{aligned} \|(\beta(f(y^n))_t - \beta(f(y^n))_s)(g_t, \cdot)\|_k &\leq C_{p,\gamma} \lambda_n \omega(s, t)^{\frac{\gamma-k}{p}}, \quad k = 1, 2, \dots, [p] - 1, \\ \|f(y_t^n) - f(y_s^n) - \beta(f(y^n))_s(g_s, g_{s,t})\| &\leq C_{p,\gamma} \lambda_n \omega(s, t)^{\frac{2}{p}}. \end{aligned}$$

Since  $y^{n+1} = \xi + \int_0^\cdot f(y^n) dx$ , by using Lemma 11, we have

$$\begin{aligned} \|(\zeta_t^{n+1} - \zeta_s^{n+1})(g_t, \cdot)\|_k &\leq C_{p,\gamma} \lambda_n \omega(s, t)^{\frac{\gamma+1-k}{p}}, \quad \forall s < t, \quad k = 1, 2, \dots, [p], \\ \|\zeta_t^{n+1}(g_t, \cdot)\| &\leq 1 + C_{p,\gamma} \lambda_n, \quad \forall t, \end{aligned}$$

which implies (6).

For the general case, we rescale the differential equation and consider  $dy = \hat{f}(y) d\hat{x}$ ,  $y_0 = \xi$ , with  $c := \lambda_{p,\gamma}^{-1} \|f\|_{\text{Lip}(\gamma)}$ ,  $\hat{f} := c^{-1}f$  and  $\hat{g} := \sum_{k=0}^{[p]} c^k \pi_k(g)$  with  $\hat{x} := \pi_1(\hat{g})$ . Then the solution path stays unchanged, and we have  $\|\hat{f}\|_{\text{Lip}(\gamma)} \leq \lambda_{p,\gamma}$ . If we denote by  $\{\beta^n\}_n$  the one-forms (as in Definition 16) associated with the Picard iterations of  $dy = \hat{f}(y) d\hat{x}$ ,  $y_0 = \xi$ , then it can be proved inductively that,

$$\zeta_s^n(g_t, v) = \beta_s^n(\hat{g}_t, \hat{v}), \quad \forall v \in \mathbb{R} \oplus \mathcal{V} \oplus \dots \oplus \mathcal{V}^{\otimes [p]} \text{ with } \hat{v} := \sum_{k=0}^{[p]} c^k \pi_k(v), \quad \forall s < t, \quad \forall n \geq 1.$$

Hence, if we can prove  $\sup_{n \geq 0} \|\beta^n\|_{\gamma+1} < \infty$  then  $\sup_{n \geq 0} \|\zeta^n\|_{\gamma+1} < \infty$ . Denote  $\hat{\omega}(s, t) := c^p \omega(s, t)$  for  $s < t$ . We divide the interval  $[0, T]$  into the union of finitely many overlapping subintervals  $\cup [s_i, t_i]$  such that  $\hat{\omega}(s_i, t_i) \leq 1$  for all  $i$ . Because these subintervals overlap, we can paste their estimates together. Indeed, by using the cocyclic property, for  $s < u < t$ ,

$$(\beta_t^n - \beta_s^n)(g_t, v) = (\beta_t^n - \beta_u^n)(g_t, v) + (\beta_u^n - \beta_s^n)(g_u, g_{u,t}v), \quad \forall v \in \mathcal{V} \oplus \dots \oplus \mathcal{V}^{\otimes [p]},$$

which implies

$$\begin{aligned} \|(\beta_t^n - \beta_s^n)(\hat{g}_t, \cdot)\|_k &\leq \|(\beta_t^n - \beta_u^n)(\hat{g}_t, \cdot)\|_k + \sum_{j=k}^{[p]} \|(\beta_u^n - \beta_s^n)(\hat{g}_u, \cdot)\|_j \\ &\leq c_1 \hat{\omega}(u, t)^{\frac{\gamma+1-k}{p}} + c_2 \sum_{j=k}^{[p]} \hat{\omega}(s, u)^{\frac{\gamma+1-j}{p}} \leq c_3 \hat{\omega}(s, t)^{\frac{\gamma+1-k}{p}}, \end{aligned}$$

where  $c_i$  may depend on  $\hat{\omega}(0, T)$ . ■

**Definition 18** With the Picard iterations  $\{y^n\}_n$  in Definition 15, we define  $z^n : [0, T] \rightarrow \mathcal{U}$ ,  $n \geq 1$ , by

$$z_t^n = y_t^n - y_t^{n-1}, \quad t \in [0, T].$$

Since  $\{y^n\}_n$  are Picard iterations which satisfy  $y^{n+1} = \xi + \int_0^\cdot f(y_u^n) dx_u$  with  $y^0 \equiv \xi$ , by using the division property of  $f$  (i.e.  $f(x) - f(y) = h(x, y)(x - y)$  for all  $x, y \in \mathcal{U}$  and  $\|h\|_{\text{Lip}(\gamma-1)} \leq C \|f\|_{\text{Lip}(\gamma)}$ ), we have the recursive expression of  $\{z^n\}_n$ :

$$z_t^{n+1} = \int_0^t h(y_u^n, y_u^{n-1}) z_u^n dx_u, \text{ with } z_t^1 = f(\xi)(x_t - x_0), \quad \forall t \in [0, T].$$

By iteration, we have

$$\begin{aligned} z_t^{n+1} &= \int \dots \int_{0 < u_1 < \dots < u_n < t} h(y_{u_n}^n, y_{u_n}^{n-1}) \dots h(y_{u_1}^1, y_{u_1}^0) z_{u_1}^1 dx_{u_1} \dots dx_{u_n} \\ &= \int \dots \int_{0 < u_0 < u_1 < \dots < u_n < t} h(y_{u_n}^n, y_{u_n}^{n-1}) \dots h(y_{u_1}^1, y_{u_1}^0) f(\xi) dx_{u_0} dx_{u_1} \dots dx_{u_n}, \quad \forall t \in [0, T]. \end{aligned}$$

Then when  $n \geq [p]$ , the increment of  $z^n$  on a small interval  $[s, t]$  can be approximated by a linear combination of  $[p]$  time-varying cocyclic one-forms, and the "coefficients" of the cocyclic one-forms are in the form of high-ordered iterated integrals so decay factorially as  $n$  tends to infinity. Hence, by relying on the factorial decay of the iterated integrals, we can prove inductively that the one-forms associated with  $\{z^n\}_n$  decays factorially in operator norm, which in turn implies the convergence in operator norm of the one-forms associated with the Picard iterations.

**Definition 19** For  $\gamma > p \geq 1$ , suppose  $g \in C^{p-var}([0, T], \mathcal{G}_{[p]})$  with  $x := \pi_1(g)$  and  $f : \mathcal{U} \rightarrow L(\mathcal{V}, \mathcal{U})$  is a  $\text{Lip}(\gamma)$  function. Let  $h$  be the function obtained in the division property of  $f$  as in Lemma 1. For integers  $n \geq l \geq 0$  and  $0 \leq s \leq t \leq T$ , we define  $\eta_{s,t}^{l,n} \in L(\mathcal{U}, \mathcal{U})$ ,  $l \geq 1$ , and  $\eta_{s,t}^{0,n} \in \mathcal{U}$ , recursively by

$$\eta_{s,t}^{l,n+1} := \int_s^t h(y_u^{n+1}, y_u^n) \eta_{s,u}^{l,n} dx_u,$$

$$\text{with } \eta_{s,t}^{l,l} := \int_s^t h(y_u^l, y_u^{l-1}) dx_u, \quad l \geq 1, \text{ and } \eta_{s,t}^{0,0} := f(\xi)(x_t - x_s).$$

The integrals are well-defined based on Lemma 11 and inductive arguments. In particular, we have

$$z_t^{n+1} = \eta_{0,t}^{0,n}, \quad \forall t \in [0, T].$$

We define  $\eta_{s,t}^{l,n}$  for general  $l$  and  $s$  to make the induction work.

Then we define the integrable one-form  $\beta_{s,\cdot}^{l,n}$  associated with the dominated path  $\eta_{s,\cdot}^{l,n}$  and prove that  $\beta_{s,\cdot}^{l,n}$  decay factorially in operator norm as  $(n-l)$  tends to infinity.

For  $\sigma_1, \sigma_2 \in \mathcal{P}_{[p]}$ ,  $|\sigma_1| + |\sigma_2| \leq [p]$ , we denote by  $\sigma_1 * \sigma_2$  the continuous linear mapping from  $\mathcal{V}^{\otimes(|\sigma_1|+|\sigma_2|)}$  to  $\mathcal{V}^{\otimes|\sigma_1|} \otimes \mathcal{V}^{\otimes|\sigma_2|}$  satisfying  $(\sigma_1 * \sigma_2)(a) = \sigma_1(a) \otimes \sigma_2(a)$  for all  $a \in \mathcal{G}_{[p]}$  (see [9] for more details).

**Definition 20** With  $\eta_{s,t}^{l,n}$  in Definition 19, for integers  $n \geq l \geq 0$  and  $s \in [0, T]$ , we define the integrable one-form  $\beta_{s,\cdot}^{l,n} : [s, T] \rightarrow B(\mathcal{G}_{[p]}, L(\mathcal{U}, \mathcal{U}))$ ,  $l \geq 1$ , and  $\beta_{s,\cdot}^{0,n} : [s, T] \rightarrow B(\mathcal{G}_{[p]}, \mathcal{U})$  (associated with  $\eta_{s,\cdot}^{l,n}$  and  $\eta_{s,\cdot}^{0,n}$  respectively) recursively by, for  $t \in (s, T]$  and  $a, b \in \mathcal{G}_{[p]}$ ,

$$\begin{aligned} \beta_{s,t}^{l,n+1}(a, b) &= \beta_{t,t}^{n+1,n+1}(a, b) \eta_{s,t}^{l,n} + h(y_t^{n+1}, y_t^n) \beta_{s,t}^{l,n}(g_t, \cdot) \pi_1(\cdot) \mathcal{I}'(g_t^{-1}a(b-1)) \\ &\quad + \beta(h(y_t^{n+1}, y_t^n))_t(g_t, \cdot) \beta_{s,t}^{l,n}(g_t, \cdot) \sum_{\sigma_i \in \mathcal{P}_{[p]}, |\sigma_1|+|\sigma_2| \leq [p]} (\sigma_1 * \sigma_2)(\cdot) \pi_1(\cdot) \mathcal{I}'(g_t^{-1}a(b-1)), \\ \beta_{s,t}^{l,l}(a, b) &= h(y_t^l, y_t^{l-1}) \pi_1(g_t^{-1}a(b-1)) + \beta(h(y_t^l, y_t^{l-1}))_t(g_t, \cdot) \pi_1(\cdot) \mathcal{I}'(g_t^{-1}a(b-1)), \quad l \geq 1, \\ \beta_{s,t}^{0,0}(a, b) &= f(\xi) \pi_1(g_t^{-1}a(b-1)), \end{aligned}$$

where  $\beta(h(y_t^{n+1}, y_t^n))$  is defined from  $(y_t^{n+1}, y_t^n)_t = (\xi, \xi) + \int_0^t (\zeta_u^{n+1}, \zeta_u^n)(g_u) dg_u$  as in Definition 12.

The notation in the definition of  $\beta_{s,\cdot}^{l,n+1}$  may need some explanations. For  $k = 1, \dots, [p] - 1$  and  $v \in \mathcal{V}^{\otimes(k+1)}$ , we have  $\mathcal{I}'(v) \in \mathcal{V}^{\otimes k} \otimes \mathcal{V}$ . Since  $\sigma_1 * \sigma_2 : \mathcal{V}^{\otimes(|\sigma_1|+|\sigma_2|)} \rightarrow \mathcal{V}^{\otimes|\sigma_1|} \otimes \mathcal{V}^{\otimes|\sigma_2|}$  and  $\pi_1 : \mathcal{V} \rightarrow \mathcal{V}$ , we have  $(\sigma_1 * \sigma_2)(\cdot) \pi_1(\cdot) \mathcal{I}'(v) \in \mathcal{V}^{\otimes|\sigma_1|} \otimes \mathcal{V}^{\otimes|\sigma_2|} \otimes \mathcal{V}$  for any  $v \in \mathcal{V}^{\otimes(|\sigma_1|+|\sigma_2|+1)}$ . Then in the expression

$$\beta(h(y_t^{n+1}, y_t^n))_t(g_t, \cdot) \beta_{s,t}^{l,n}(g_t, \cdot) (\sigma_1 * \sigma_2)(\cdot) \pi_1(\cdot) \mathcal{I}'(v) \text{ for } v \in \mathcal{V}^{\otimes(|\sigma_1|+|\sigma_2|+1)},$$

we treat  $\beta(h(y_t^{n+1}, y_t^n))_t(g_t, \cdot)$  as a continuous linear mapping on  $\mathcal{V}^{\otimes|\sigma_1|}$  and treat  $\beta_{s,t}^{l,n}(g_t, \cdot)$  as a continuous linear mapping on  $\mathcal{V}^{\otimes|\sigma_2|}$ .

Based on the definition of integral in Definition 13, we have  $\eta_{s,t}^{l,n} = \int_s^t \beta_{s,u}^{l,n}(g_u) dg_u$ . In particular,

$$z_t^{n+1} = \int_0^t \beta_{0,u}^{0,n}(g_u) dg_u, \quad \forall t \in [0, T].$$

**Lemma 21** Suppose  $g \in C^{p-var}([0, T], \mathcal{G}_{[p]})$  is controlled by  $\omega$ , and  $f : \mathcal{U} \rightarrow L(\mathcal{V}, \mathcal{U})$  is a  $\text{Lip}(\gamma)$  function for some  $\gamma \in (p, [p] + 1]$ . Then there exist a constant  $C = C(p, \gamma, \|f\|_{\text{Lip}(\gamma)}, \omega(0, T))$  such that

$$\|\beta_{s,\cdot}^{l,n}\|_\gamma \leq \frac{C^{n-[p]-l}}{\left(\frac{n-[p]-l}{p}\right)!}, \quad \forall n \geq l + [p] + 1, \quad \forall l \geq 0, \quad \forall s \in [0, T], \quad (7)$$

where  $\beta_{s,\cdot}^{l,n}$  denotes  $t \mapsto \beta_{s,t}^{l,n}$  introduced in Definition 20 for  $t \in [s, T]$ .

**Proof.** The constants in this proof may depend on  $p, \gamma, \|f\|_{\text{Lip}(\gamma)}$  and  $\omega(0, T)$ .

Firstly, we prove that, for integers  $n \geq l \geq 0$  and  $0 \leq s \leq u \leq t \leq T$ ,

$$\beta_{s,t}^{l,n}(g_t, v) = \beta_{u,t}^{l,n}(g_t, v) + \sum_{j=l+1}^n \beta_{u,t}^{j,n}(g_t, v) \eta_{s,u}^{l,j-1}, \forall v \in \mathcal{V} \oplus \dots \oplus \mathcal{V}^{\otimes [p]}. \quad (8)$$

The equality holds when  $n = l$  based on the definition of  $\beta_{s,t}^{l,l}$ . Suppose it holds when  $n - l \leq s$ . Then by combining the definition of  $\beta_{s,t}^{l,n+1}$  in Definition 20 with the inductive hypothesis (8) and by using  $\eta_{s,t}^{l,n} = \sum_{j=l+1}^n \eta_{u,t}^{j,n} \eta_{s,u}^{l,j-1} + \eta_{s,u}^{l,n} + \eta_{u,t}^{l,n}$ , it can be proved that (8) holds when  $n - l = s + 1$ .

Without loss of generality we assume  $\gamma \in (p, [p] + 1]$ . Based on Lemma 17,  $\sup_{n \geq 0} \|\zeta^n\|_{[p]+1} < \infty$ . Then since  $\|h\|_{\text{Lip}(\gamma-1)} \leq C \|f\|_{\text{Lip}(\gamma)}$ , by using Lemma 11, it can be proved inductively that, for some  $K_0 \geq 1$ ,

$$\sup_{l \geq 0} \max_{n=l, \dots, l+[p]} \left\| \beta_{s,\cdot}^{l,n} \right\|_{\gamma} \leq K_0. \quad (9)$$

Then combined with  $\eta_{s,t}^{l,n} = \int_s^t \beta_{s,u}^{l,n}(g_u) dg_u$ , we have, for some constant  $M_0 > 0$ ,

$$\left\| \eta_{s,t}^{l,n} \right\| \leq M_0 \omega(s, t)^{\frac{n-l+1}{p}}, \forall s < t, n - l + 1 = 1, 2, \dots, [p], \forall l \geq 0.$$

Since

$$\eta_{s,t}^{l,n} = \sum_{j=l+1}^n \eta_{u,t}^{j,n} \eta_{s,u}^{l,j-1} + \eta_{s,u}^{l,n} + \eta_{u,t}^{l,n}, \forall 0 \leq s < u < t \leq T, \forall n \geq l \geq 0,$$

by following similar proof as the factorial decay of the signature of a rough path as in Theorem 3.7 [8], we have that, for  $\beta = 3p$  and some constant  $M \geq 1$ , (we choose  $\beta = 3p$  to make the induction work)

$$\left\| \eta_{s,t}^{l,n} \right\| \leq \frac{M^{\frac{n-l+1}{p}} \omega(s, t)^{\frac{n-l+1}{p}}}{\beta \left( \frac{n-l+1}{p} \right)!}, \forall s < t, \forall n \geq l \geq 0. \quad (10)$$

Then we prove by induction on  $n - l$  that, for some constants  $K \geq 1$  and  $C \geq 1$  (we will chose them in the inductive step),

$$\left\| \left( \beta_{s,t}^{l,n} - \beta_{s,u}^{l,n} \right) (g_t, \cdot) \right\|_k \leq K \frac{C^{\frac{n-[p]-l}{p}} \omega(s, t)^{\frac{n-[p]-l}{p}}}{\beta \left( \frac{n-[p]-l}{p} \right)!} \omega(u, t)^{\frac{\gamma-k}{p}}, \forall s < u < t, \forall n \geq l + [p], \forall l \geq 0, \quad (11)$$

which holds when  $n - l = [p]$  with  $K = K_0 \beta$  based on (9). Suppose (11) holds when  $n - l = [p], \dots, s$  for some  $s \geq [p]$ . Then when  $n - l = s + 1$  (so  $n - l \geq [p] + 1$ ), for  $s < u < t$ , based on (8), we have, for any  $v \in \mathcal{V} \oplus \dots \oplus \mathcal{V}^{\otimes [p]}$ ,

$$\begin{aligned} & \left( \beta_{s,t}^{l,n} - \beta_{s,u}^{l,n} \right) (g_t, v) \\ &= \left( \beta_{u,t}^{l,n} - \beta_{u,u}^{l,n} \right) (g_t, v) + \sum_{j=n-[p]}^n \left( \beta_{u,t}^{j,n} - \beta_{u,u}^{j,n} \right) (g_t, v) \eta_{s,u}^{l,j-1} + \sum_{j=l+1}^{n-[p]-1} \left( \beta_{u,t}^{j,n} - \beta_{u,u}^{j,n} \right) (g_t, v) \eta_{s,u}^{l,j-1} \\ &=: I(v) + II(v) + III(v). \end{aligned} \quad (12)$$

For  $I(v)$ , by using (8), we have

$$\left( \beta_{u,t}^{l,n} - \beta_{u,u}^{l,n} \right) (g_t, v) = \left( \beta_{t,t}^{l,n} - \beta_{u,u}^{l,n} \right) (g_t, v) + \sum_{j=l+1}^n \beta_{t,t}^{j,n}(g_t, v) \eta_{u,t}^{l,j-1} = \sum_{j=n-[p]+1}^n \beta_{t,t}^{j,n}(g_t, v) \eta_{u,t}^{l,j-1},$$

where we used  $\beta_{t,t}^{j,n} \equiv 0$  for  $n \geq [p] + j$  which can be proved inductively based on the definition of  $\beta_{s,t}^{l,n}$  in Definition 20. Hence, for  $k = 1, \dots, [p]$ , by using that  $\left\| \beta_{t,t}^{j,n}(g_t, \cdot) \right\|_k = 0, j \leq n - k$ , and the factorial decay of  $\eta_{s,t}^{l,n}$  in (10), we have, for some  $C_0 \geq 1$ , (since  $\gamma \leq [p] + 1$ )

$$\begin{aligned} \|I(\cdot)\|_k &= \left\| \sum_{j=n-[p]+1}^n \beta_{t,t}^{j,n}(g_t, \cdot) \eta_{u,t}^{l,j-1} \right\|_k \\ &\leq \sum_{j=n-k+1}^n \left\| \beta_{t,t}^{j,n}(g_t, \cdot) \right\|_k \frac{M^{\frac{j-l}{p}} \omega(u, t)^{\frac{j-l}{p}}}{\beta \left( \frac{j-l}{p} \right)!} \leq K_0 C_0 \frac{M^{\frac{n-[p]-l}{p}} \omega(u, t)^{\frac{n-[p]-l}{p}}}{\beta \left( \frac{n-[p]-l}{p} \right)!} \omega(u, t)^{\frac{\gamma-k}{p}}. \end{aligned} \quad (13)$$

For  $II(v)$ , by using (9) and (10), we have

$$\begin{aligned} \|II(\cdot)\|_k &= \left\| \sum_{j=n-[p]}^n \left( \beta_{u,t}^{j,n} - \beta_{u,u}^{j,n} \right) (g_t, \cdot) \eta_{s,u}^{l,j-1} \right\|_k \\ &\leq \sum_{j=n-[p]}^n \left\| \left( \beta_{u,t}^{j,n} - \beta_{u,u}^{j,n} \right) (g_t, \cdot) \right\|_k \frac{M^{\frac{j-l}{p}} \omega(s, u)^{\frac{j-l}{p}}}{\beta\left(\frac{j-l}{p}\right)!} \leq K_0 C_0 \frac{M^{\frac{n-[p]-l}{p}} \omega(s, u)^{\frac{n-[p]-l}{p}}}{\beta\left(\frac{n-[p]-l}{p}\right)!} \omega(u, t)^{\frac{\gamma-k}{p}}. \end{aligned} \quad (14)$$

For  $III(v)$ , since  $[p] < n - j \leq n - l - 1 = s$  when  $j = l + 1, \dots, n - [p] - 1$ , by using the inductive hypothesis (11) and neo-classical inequality [8, 5], we have

$$\begin{aligned} \|III(\cdot)\|_k &= \left\| \sum_{j=l+1}^{n-[p]-1} \left( \beta_{u,t}^{j,n} - \beta_{u,u}^{j,n} \right) (g_t, \cdot) \eta_{s,u}^{l,j-1} \right\|_k \\ &\leq \sum_{j=l+1}^{n-[p]-1} K \frac{C^{\frac{n-[p]-j}{p}} \omega(u, t)^{\frac{n-[p]-j}{p}}}{\beta\left(\frac{n-[p]-j}{p}\right)!} \frac{M^{\frac{j-l}{p}} \omega(s, u)^{\frac{j-l}{p}}}{\beta\left(\frac{j-l}{p}\right)!} \omega(u, t)^{\frac{\gamma-k}{p}} \\ &\leq K \frac{p(C \vee M)^{\frac{n-[p]-l}{p}} \omega(s, t)^{\frac{n-[p]-l}{p}}}{\beta\left(\frac{n-[p]-l}{p}\right)!} \omega(u, t)^{\frac{\gamma-k}{p}}. \end{aligned} \quad (15)$$

Hence, based on (12), (13), (14) and (15), since  $n - [p] - l \geq 1$  and  $\beta = 3p$ , by choosing  $K = K_0(C_0 \vee \beta)$  ( $\beta$  to take into account of  $n = l + [p]$ ) and  $C = 3M$ , we have (11) holds when  $n - l = s + 1$ , and the induction is complete.

On the other hand, when  $n - l \geq [p]$ , based on (8) and by using  $\beta_{t,t}^{j,n} \equiv 0$  for  $n \geq [p] + j$ , we have

$$\beta_{s,t}^{l,n}(g_t, \cdot) = \beta_{t,t}^{l,n}(g_t, \cdot) + \sum_{j=l+1}^n \beta_{t,t}^{j,n}(g_t, \cdot) \eta_{s,t}^{l,j-1} = \sum_{j=n-[p]+1}^n \beta_{t,t}^{j,n}(g_t, \cdot) \eta_{s,t}^{l,j-1}.$$

Hence, by using the factorial decay of  $\eta_{s,t}^{l,n}$  in (10), we have

$$\left\| \beta_{s,t}^{l,n}(g_t, \cdot) \right\| \leq \sum_{j=n-[p]+1}^n \left\| \beta_{t,t}^{j,n}(g_t, \cdot) \right\|_k \left\| \eta_{s,t}^{l,j-1} \right\| \leq K_0 C_0 \frac{M^{\frac{n-[p]-l}{p}} \omega(s, t)^{\frac{n-[p]-l}{p}}}{\left(\frac{n-[p]-l}{p}\right)!}. \quad (16)$$

Then for  $s \in [0, T]$  since

$$\left\| \beta_{s,\cdot}^{l,n} \right\|_{\gamma} := \sup_{s \leq t \leq T} \left\| \beta_{s,t}^{l,n}(g_t, \cdot) \right\| + \max_{k=1, \dots, [p]} \sup_{s \leq u \leq t \leq T} \omega(u, t)^{-(\gamma - \frac{k}{p})} \left\| \left( \beta_{s,t}^{l,n} - \beta_{s,u}^{l,n} \right) (g_t, \cdot) \right\|_k,$$

we have the lemma holds based on (11) and (16). ■

**Theorem 22 (Existence, Uniqueness and Continuity of the Solution)** For  $[p] + 1 \geq \gamma > p \geq 1$ , suppose  $g \in C^{p-var}([0, T], \mathcal{G}_{[p]})$  is controlled by  $\omega$ ,  $f : \mathcal{U} \rightarrow L(\mathcal{V}, \mathcal{U})$  is a  $\text{Lip}(\gamma)$  function and  $\xi \in \mathcal{U}$ . Then the Picard iterations  $\{y^n\}_{n=0}^{\infty}$  in Definition 15 converge uniformly on  $[0, T]$  to the unique solution to the rough differential equation

$$dy = f(y) dx, \quad y_0 = \xi,$$

and the solution is continuous with respect to  $g$  in  $p$ -variation norm. Moreover, there exist integrable one-forms  $\beta^n : [0, T] \rightarrow B(\mathcal{G}_{[p]}, \mathcal{U})$ ,  $n \geq 0$ , and a constant  $C = C(p, \gamma, \|f\|_{\text{Lip}(\gamma)}, \omega(0, T)) > 0$  such that

$$\begin{aligned} y_t^n &= \xi + \int_0^t \beta_u^n(g_u) dg_u, \quad \forall t \in [0, T], \\ \text{and } \left\| \beta^{n+1} - \beta^n \right\|_{\gamma} &\leq \frac{C^{n-[p]}}{\left(\frac{n-[p]}{p}\right)!}, \quad \forall n \geq [p] + 1. \end{aligned} \quad (17)$$

There are some remarks.

1. In proving the convergence of the Picard iterations, we proved the convergence in operator norm of their associated one-forms. In particular, we proved that the one-form associated with the difference between the  $n$ th and  $(n+1)$ th Picard iterations decays factorially on  $[0, T]$  as  $n$  tends to infinity.
2. Let  $\rho : [0, T] \rightarrow \mathcal{W}$  be a path dominated by  $g$ . Then the integral of  $\rho$  against  $y^n$  is well defined:

$$\int_0^t \rho_u \otimes dy_u^n = \int_0^t \rho_u \otimes f(y_u^{n-1}) dx_u, \forall t \in [0, T].$$

In particular, since  $y^n$  is a dominated path, there exists a canonical enhancement of  $y^n$  to a group-valued path, which could take values in nilpotent Lie group or Butcher group.

3. When treated as a Banach space-valued path, the group-valued enhancement is again a dominated path. Since the one-form associated with the enhancement is continuous with respect to the one-form associated with the base dominated path, the one-forms of the enhancement of  $y^n$  also converge in operator norm, which implies the uniform convergence of the group-valued enhancements.
4. When  $f$  is  $\text{Lip}(\gamma)$  for  $\gamma > p-1$ , the one-forms associated with the Picard iterations are uniformly bounded. When the dimension is finite, based on Arzelà-Ascoli theorem, there exists a subsequence of the one-forms which converges, so the associated paths (and their enhancements) converge to a solution.
5. When  $f$  is locally Lipschitz and the dimension is finite, the solution exists (uniquely) up to explosion. Indeed, by Whitney's extension theorem, the restriction of  $f$  to any compact set can be extended to a global Lipschitz function without increasing its Lipschitz norm, so the solution exists up to exit time of that compact set. For similar reason, when  $f$  is locally  $\text{Lip}(\gamma)$  for  $\gamma > p$ , any two solutions must agree on any compact set, so the solution exists uniquely up to explosion.

**Proof.** Suppose  $\{y^n\}_n$  are the Picard iterations in Definition 15. Since  $z^{i+1} = y^{i+1} - y^i$  and  $\beta^{0,i}$  is the integrable one-form associated with  $z^{i+1}$ , if we define  $\beta^n : [0, T] \rightarrow B(\mathcal{G}_{[p]}, \mathcal{U})$ ,  $n \geq 1$ , by

$$\beta_s^n(a, b) = \sum_{i=0}^{n-1} \beta_s^{0,i}(a, b), \forall a, b \in \mathcal{G}_{[p]}, \forall s \in [0, T],$$

then  $\beta^n$  is integrable and

$$y_t^n = \xi + \int_0^t \beta_u^n(g_u) dg_u, \forall t \in [0, T], \forall n \geq 1.$$

(Since  $y^0 \equiv \xi$ , we set  $\beta^0 \equiv 0$  so  $y^0 = \xi + \int_0^\cdot \beta^0(g) dg$ .) Based on Lemma 21, we have (17) holds and  $\beta^n$  converge in operator norm as  $n$  tends to infinity (denote the limit by  $\beta$ ), so  $y^n = \xi + \int_0^\cdot \beta^n(g) dg$  converge uniformly to  $y := \xi + \int_0^\cdot \beta(g) dg$ . Moreover, by using the division property of  $f$  (i.e.  $f(x) - f(y) = h(x, y)(x - y)$  for all  $x, y$  in  $\mathcal{U}$  and  $\|h\|_{\text{Lip}(\gamma-1)} \leq C\|f\|_{\text{Lip}(\gamma)}$ ), we have

$$\begin{aligned} y_t^{n+1} - y_t^n &= z_t^{n+1} = \int_0^t h(y_u^n, y_u^{n-1}) z_u^n dx_u \\ &= \int_0^t h(y_u^n, y_u^{n-1}) (y_u^n - y_u^{n-1}) dx_u = \int_0^t (f(y_u^n) - f(y_u^{n-1})) dx_u, \forall t \in [0, T]. \end{aligned}$$

Hence,

$$y_t^{n+1} = \xi + \int_0^t f(y_u^n) dx_u, \forall t \in [0, T], \forall n \geq 0, \text{ with } y^0 \equiv \xi.$$

Since both  $y^n$  and  $y^{n+1}$  are dominated paths and their associated one-forms converge to  $\beta$  as  $n$  tends to infinity, by letting  $n \rightarrow \infty$  on both sides, we have  $\beta$  is the fixed point of the mapping  $\beta \mapsto \hat{\beta}$  where  $\hat{\beta}$  is the one-form associated with the dominated path  $t \mapsto \int_0^t f(y) dx$ . Hence,  $y$  is a dominated path satisfying the integral equation and  $y$  is a solution.

Then we prove that the solution is unique. Suppose  $\hat{y}$  is another solution. By using the division property of  $f$ , we have

$$y_t - \hat{y}_t = \int_0^t (f(y_u) - f(\hat{y}_u)) dx_u = \int_0^t h(y_u, \hat{y}_u) (y_u - \hat{y}_u) dx_u, \forall t \in [0, T].$$

By iterating this process, we have, for any integer  $n \geq 1$ ,

$$y_t - \hat{y}_t = \int \cdots \int_{0 < u_1 < \cdots < u_n < t} h(y_{u_n}, \hat{y}_{u_n}) \cdots h(y_{u_1}, \hat{y}_{u_1}) (y_{u_1} - \hat{y}_{u_1}) dx_{u_1} \cdots dx_{u_n}, \forall t \in [0, T].$$

Since  $(y, \hat{y})$  is a dominated path and  $h$  is a  $\text{Lip}(\gamma - 1)$  function, we can define based on Lemma 11 the dominated paths  $\rho^n : [0, T] \rightarrow L(\mathcal{U}, \mathcal{U})$ ,  $n \geq 1$ , recursively by

$$\rho_t^{n+1} = \int_0^t h(y_u, \hat{y}_u) \rho_u^n dx_u \text{ with } \rho_t^1 = \int_0^t h(y_u, \hat{y}_u) dx_u, \forall t \in [0, T],$$

and we have

$$y_t - \hat{y}_t = \int_0^t \rho_u^n (y_u - \hat{y}_u) dx_u, \forall t \in [0, T], \forall n \geq 1.$$

Then by following similar proof to that of Lemma 21, the one-form associated with  $\rho^n$  decays factorially. Since  $y - \hat{y}$  is another dominated path, the one-form associated with the dominated path  $\int_0^\cdot \rho_u^n (y_u - \hat{y}_u) dx_u$  also decays factorially, which implies that  $y = \hat{y}$ .

It is clear that for any integer  $n \geq 1$ , the mapping  $g \mapsto \beta^n$  is continuous. Suppose  $g^m \rightarrow g$  in  $p$ -variation norm, then by uniform convergence of the mapping  $\beta^n \mapsto \beta$  with respect to the  $p$ -variation of  $g$  (based on Lemma 21), we have  $g \mapsto \beta$  is continuous, which implies that the mapping  $g \mapsto y$  is continuous with respect to  $g$  in  $p$ -variation norm. ■

The Oxford-Man Institute, University of Oxford

## References

- [1] J. C. Butcher. An algebraic theory of integration methods. *Mathematics of Computation*, 26(117):79–106, 1972.
- [2] A. Connes and D. Kreimer. Hopf algebras, renormalization and noncommutative geometry. In *Quantum field theory: perspective and prospective*, pages 59–109. Springer, 1999.
- [3] M. Gubinelli. Controlling rough paths. *Journal of Functional Analysis*, 216(1):86–140, 2004.
- [4] M. Gubinelli. Ramification of rough paths. *Journal of Differential Equations*, 248(4):693–721, 2010.
- [5] K. Hara and M. Hino. Fractional order taylor’s series and the neo-classical inequality. *Bulletin of the London Mathematical Society*, 42(3):467–477, 2010.
- [6] T. Lyons. Differential equations driven by rough signals. i. an extension of an inequality of lc young. *Math. Res. Lett*, 1(4):451–464, 1994.
- [7] T. J. Lyons. Differential equations driven by rough signals. *Rev. Mat. Iberoamericana*, 14(2), 1998.
- [8] T. J. Lyons, M. Caruana, and T. Lévy. *Differential equations driven by rough paths*. Springer, 2007.
- [9] T. J. Lyons and D. Yang. Integration of time-varying cocyclic one-forms against rough paths. *arXiv preprint arXiv:1408.2785*, 2014.
- [10] C. Reutenauer. Free lie algebras. *Handbook of algebra*, 3:887–903, 2003.
- [11] L. Schwartz. La convergence de la série de picard pour les eds (equations différentielles stochastiques). In *Séminaire de Probabilités XXIII*, pages 343–354. Springer, 1989.
- [12] E. M. Stein. *Singular integrals and differentiability properties of functions*, volume 2. Princeton university press, 1970.
- [13] L. C. Young. An inequality of the hölder type, connected with stieltjes integration. *Acta Mathematica*, 67(1):251–282, 1936.